

## SEQUENCE LISTING

<110> SHIBATA, Takashi

ICHIKAWA, Chiyo

MATSUURA, Mitsutaka

NOGUCHI, Yuji

SAITO, Yoshimasa

YAMASHITA, Michio

TAKATA, Yoko

<120> SORBITOL DEHYDROGENASE, GENE ENCODING THE SAME AND USE THEREOF

312  
<130> 213930US0PCT

<140> 09/926,163

<141> 2001-09-17

<150> PCT/JP00/01608

<151> 2000-03-16

<150> JP11/72810

<151> 1999-03-17

<150> JP11/224679

<151> 1999-08-06

<160> 20

<170> PatentIn version 3.1

<210> 1

<211> 4115

<212> DNA

<213> Gluconobacter oxydans

Q12  
<220>

<221> CDS

<222> (537) .. (1994)

<223>

<400> 1

aagcttgcat	gcctgcaggt	cgactctaga	ggatccgggt	ttggcagcgc	tccttagatt	60
gatgcggcgt	ctggtgaccg	acatgatgct	ggtggcacgt	gccattgcga	cggggcgtgc	120
gaccgggaac	acaggcctgc	tgcctttgta	caaggggctg	agtcatgcgc	tgcgtgggct	180
ggcacatagt	tgcgaagagc	agttgcgcgc	aaagcagaac	cagcatgaac	agcagtcgca	240
agacgaggaa	atcctcggcc	tcctaccgcg	attggaagag	cagaccgcgc	ctgagatgcg	300
ttttgtgatg	tcctgtttcc	gcgaggatct	cgaacgggct	gttgggggtgc	tcatgcgttc	360
tgatgcgagt	gccgcaaaag	gtctctgaac	aggacgtccc	gcggagggca	gtcagaggtc	420
gaaatggctc	ctggtgaaac	cgtcattcgg	tttttacggt	gtttcggggc	tatgatggca	480
catgcccggc	cttgtcgggtc	cccgtcagcg	accggcccga	aaccacggag	aattcc atg	539
					Met	
					1	
att acg cgc	gaa acc ctt	aag tct ctt	cct gcc aat	gtc cag gct	ccc	587
Ile Thr Arg	Glu Thr Leu	Lys Ser Leu	Pro Ala Asn	Val Gln Ala	Pro	
	5		10		15	
ccc tat gac	atc gac ggg	atc aag cct	ggg atc gtg	cat ttc ggt	gta	635
Pro Tyr Asp	Ile Asp Gly	Ile Lys Pro	Gly Ile Val	His Phe Gly	Val	

20	25	30	
ggt aac ttt ttt cga gcc cat gag gcg ttc tac gtc gag cag att ctt Gly Asn Phe Phe Arg Ala His Glu Ala Phe Tyr Val Glu Gln Ile Leu 35 40 45	683		
gaa cac gct ccg gac tgg gcg att gtt ggt gtt ggc ctg acg ggc agt Glu His Ala Pro Asp Trp Ala Ile Val Gly Val Gly Leu Thr Gly Ser 50 55 60 65	731		
gac cgt tca aag aaa aaa gcc gag gaa ttc aag gcc cag gac tgc ctg Asp Arg Ser Lys Lys Lys Ala Glu Glu Phe Lys Ala Gln Asp Cys Leu 70 75 80	779		
tat tcc ctg acc gag acg gct ccg tcc ggc aag agc acg gtg cgc gtc Tyr Ser Leu Thr Glu Thr Ala Pro Ser Gly Lys Ser Thr Val Arg Val 85 90 95	827		
atg ggc gcg ctg cgt gac tat ctg ctt gcc ccg gcc gat ccg gaa gcc Met Gly Ala Leu Arg Asp Tyr Leu Leu Ala Pro Ala Asp Pro Glu Ala 100 105 110	875		
gtg ctg aag cat ctt gtt gat ccg gcc atc cgc atc gtt tcc atg acg Val Leu Lys His Leu Val Asp Pro Ala Ile Arg Ile Val Ser Met Thr 115 120 125	923		
atc acg gaa ggc ggc tac aac atc aac gag acg acc ggt gcg ttc gat Ile Thr Glu Gly Gly Tyr Asn Ile Asn Glu Thr Thr Gly Ala Phe Asp 130 135 140 145	971		
ctg gag aat gcg gca gta aag gcc gac ctc aag aac ccg gaa aag ccg Leu Glu Asn Ala Ala Val Lys Ala Asp Leu Lys Asn Pro Glu Lys Pro 150 155 160	1019		
tct acc gtt ttc ggt tac gtg gtc gag gcc ctg cgt cgt cgt tgg gat Ser Thr Val Phe Gly Tyr Val Val Glu Ala Leu Arg Arg Arg Trp Asp 165 170 175	1067		
gcc ggt ggt aag gca ttt acg gtc atg tcc tgt gat aac ctg cgt cat Ala Gly Gly Lys Ala Phe Thr Val Met Ser Cys Asp Asn Leu Arg His 180 185 190	1115		
aac ggc aat gtc gcc cgc aag gcc ttc ctc ggc tat gcg aag gcg cgc Asn Gly Asn Val Ala Arg Lys Ala Phe Leu Gly Tyr Ala Lys Ala Arg 195 200 205	1163		
gat ccg gag ttg gcg aag tgg att gag gaa aac gcg acc ttc ccg aac Asp Pro Glu Leu Ala Lys Trp Ile Glu Glu Asn Ala Thr Phe Pro Asn 210 215 220 225	1211		
gga atg gtt gat cgc atc acc ccg acc gtt tcg gcg gaa atc gcc aag Gly Met Val Asp Arg Ile Thr Pro Thr Val Ser Ala Glu Ile Ala Lys	1259		

230						235						240						
aag	ctc	aac	gcg	gcc	agt	ggg	ctg	gat	gac	gac	ctg	ccg	ctg	gtg	gcc			1307
Lys	Leu	Asn	Ala	Ala	Ser	Gly	Leu	Asp	Asp	Asp	Leu	Pro	Leu	Val	Ala			
			245					250					255					
gag	gat	ttc	cat	cag	tgg	gtg	ctg	gaa	gac	cag	ttt	gcg	gat	ggc	cgt			1355
Glu	Asp	Phe	His	Gln	Trp	Val	Leu	Glu	Asp	Gln	Phe	Ala	Asp	Gly	Arg			
		260					265					270						
ccg	ccg	ctt	gaa	aaa	gcc	ggc	gtg	cag	atg	gtc	ggg	gac	gtg	acg	gac			1403
Pro	Pro	Leu	Glu	Lys	Ala	Gly	Val	Gln	Met	Val	Gly	Asp	Val	Thr	Asp			
	275					280					285							
tgg	gag	tac	gtc	aag	atc	cga	atg	ctc	aat	gca	ggg	cat	gtc	atg	ctc			1451
Trp	Glu	Tyr	Val	Lys	Ile	Arg	Met	Leu	Asn	Ala	Gly	His	Val	Met	Leu			
290					295					300					305			
tgc	ttc	cca	ggc	att	ctg	gtc	ggc	tat	gag	aat	gtg	gat	gac	gcc	att			1499
Cys	Phe	Pro	Gly	Ile	Leu	Val	Gly	Tyr	Glu	Asn	Val	Asp	Asp	Ala	Ile			
				310					315					320				
gaa	gac	agc	gaa	ctc	ctt	ggc	aat	ctg	aag	aac	tat	ctc	aac	aag	gat			1547
Glu	Asp	Ser	Glu	Leu	Leu	Gly	Asn	Leu	Lys	Asn	Tyr	Leu	Asn	Lys	Asp			
			325					330					335					
gtc	atc	ccg	acc	ctg	aag	gcg	cct	tca	ggc	atg	acg	ctc	gaa	ggc	tat			1595
Val	Ile	Pro	Thr	Leu	Lys	Ala	Pro	Ser	Gly	Met	Thr	Leu	Glu	Gly	Tyr			
		340					345					350						
cgg	gac	agc	gtc	atc	agc	cgt	ttc	tcc	aac	aag	gcg	atg	tcg	gac	cag			1643
Arg	Asp	Ser	Val	Ile	Ser	Arg	Phe	Ser	Asn	Lys	Ala	Met	Ser	Asp	Gln			
	355					360					365							
acg	ctc	cgg	att	gct	agc	gat	ggc	tgt	tcc	aag	gtt	cag	gtg	ttc	tgg			1691
Thr	Leu	Arg	Ile	Ala	Ser	Asp	Gly	Cys	Ser	Lys	Val	Gln	Val	Phe	Trp			
370					375					380					385			
acg	gaa	acc	gtg	cgt	cgg	gcg	atc	gaa	gac	aag	cgg	gac	ctg	tca	cgt			1739
Thr	Glu	Thr	Val	Arg	Arg	Ala	Ile	Glu	Asp	Lys	Arg	Asp	Leu	Ser	Arg			
				390					395					400				
ata	gcg	ttc	gga	att	gca	tcc	tat	ctc	gaa	atg	ctg	cgt	ggc	cgc	gac			1787
Ile	Ala	Phe	Gly	Ile	Ala	Ser	Tyr	Leu	Glu	Met	Leu	Arg	Gly	Arg	Asp			
			405					410				415						
gag	aag	ggc	ggg	acg	tat	gaa	tcg	tcc	gag	ccg	act	tat	ggc	gac	gcc			1835
Glu	Lys	Gly	Gly	Thr	Tyr	Glu	Ser	Ser	Glu	Pro	Thr	Tyr	Gly	Asp	Ala			
		420					425					430						
gaa	tgg	aag	ttg	gcc	aag	gcg	gac	gac	ttc	gaa	agc	tct	ctg	aag	ctc			1883
Glu	Trp	Lys	Leu	Ala	Lys	Ala	Asp	Asp	Phe	Glu	Ser	Ser	Leu	Lys	Leu			

435	440	445	
ccg gcg ttc gat ggg tgg cgc gat ctg gat acg tcc gaa ctg gat caa			1931
Pro Ala Phe Asp Gly Trp Arg Asp Leu Asp Thr Ser Glu Leu Asp Gln			
450	455	460	465
aag gtc atc gtg ctg cgg aag atc atc cgc gaa aag ggc gta aaa gcc			1979
Lys Val Ile Val Leu Arg Lys Ile Ile Arg Glu Lys Gly Val Lys Ala			
	470	475	480
gcc atc ccg gcc tga attcggccttt tagggtagcg actgaaacag aaaaccgcgc			2034
Ala Ile Pro Ala			
	485		
tctggaagga ggcgcggtttt ttttatgctc agatctgtcc catcaggaca aggatcacga			2094
cgaccacgat caggacaagt ccgctggagg gggagcccca tttcgaactg tacggccatg			2154
acggcagcgc accgagatca ggattacaag aaggatcagt cccatggcac atctctcttg			2214
212 ccggttgaga ctggtctgtg ttccgggtgt ctaaaaagtt tccgtagggg cgcgaaagat			2274
caaagctgtc ggtcgcgctt aatccgggtcc caagccgcat tgatgcgggc caccgcgtcc			2334
tgtgcgcggt tgcgctctgt ctctgacata ggtttctggg ccagcacgtc cggatgatgt			2394
tcgcggatca ggggtgcgcca gcgcacgcgg atttctgtgt cagttgcgct gcgggtgatg			2454
ccgagaatac gataggcatc cggctcgttt ccgctggcgg cgcgattggt gccgctttcg			2514
gcccgggtccc atgctcctgg cggcaggcca aatgccccgt gaacgcgctg cagaaaatcg			2574
atttccttcg ggtgaagctc gcggctgggg ccggcatcgg cacgggcatg acggaacagt			2634
gccgtcatga ggttctcaag cggcgccgta ttatcggcat aggccttgcc catttcgcgg			2694
gcatacatct cgaaatcgtc cgtccgggtcg cgggcgcgat cgaacagcat gccgacttcc			2754
ttggtgttat cgggggggaa ctggaagcag gtcttgaaag cgttgatttc gtgtcggttc			2814
accggcccgt cgatcttcgc cagcttcgcg cacagggcaa caaggccgat ggcgtaaagc			2874
tgatctcgtt tgcccagggc cgcagcaatc ttggcagcgc cgaaaaaggc cgcgctgttg			2934
ggatcgggac ggccattcgc gggaaagcgc tcaactccagc cgcccgttga gggcttgagt			2994
agcgaaccgt tatcggcggc atgccccagc gctgcgcca tcagtgtctc gaaaggacca			3054
ccaaccgcga agcccgcgac accaccgaac atcttgcccc agatagccat gtcatcaacc			3114
tagcacgccc gctcacagcg gcaaatgaca gatcgcaggc taggtgtagg tgctgatgcg			3174

ccaaccgccc gggcttgccg tgtggtagaa gctaggagtt acgaacttat cgctgtctca 3234  
 tgcttttgag gcgcagggtc ttctgttcgt ttcattgacgg atatttttat gccacacttg 3294  
 atccagactg ctacttcgat ccctttccgc tctgatgacg aactgatgga tcttttgatc 3354  
 aagcgtctgc caatgtggct gcagaaagtg ctgaactggg tgcgggaagc ggatcataaa 3414  
 tgggttcgga ttccggcggg cgtgctgttc atgctgggcg gcgttctgtc catcctgcct 3474  
 gttctgggtc tgtggatgct gccggtcggc gtgatgttgc ttgcgcagga tattccgttc 3534  
 ttccgtcgcc ttcagggccg cctcttgccg tggatcgaac gtcaacatcc ggattggctg 3594  
 ggctttccgg cgaaaagcgg cagaagctaa ccgttcgtct ggacgtgttt ctgaagatgt 3654  
 gtcagtgtg caacccgcag ggctgaagcc agtgggcgct ctggtggctg cgcggcatcg 3714  
 212 agagaagcca ccagagacgc aaagctctgc tggcggactg cggccatcgc gtccagtata 3774  
 gcccagaact cgggttccag tgccacggac gtccggtgtc ctgacagaga caggctgcgt 3834  
 ttgacgagat cactcattcc ggttggtttct caaggcgctt caaagcccat tgtgcggttt 3894  
 cggaacatc agggtcgga tcaactcagca gctcccgcgc agaagatata agcgacggat 3954  
 cggccgagtt gccgatcgcg atcaggacag ttacgtacga accgggtgcg tccaatccgt 4014  
 ttgaccggag agccagaaaa aaacgtccgg aatgtcgcat tatccagccg caccagttcg 4074  
 tcgagttttg gtgcaatcag ctccgggccc gctgaagct t 4115

<210> 2

<211> 485

<212> PRT

<213> Gluconobacter oxydans

<400> 2

Met Ile Thr Arg Glu Thr Leu Lys Ser Leu Pro Ala Asn Val Gln Ala  
 1 5 10 15

Pro Pro Tyr Asp Ile Asp Gly Ile Lys Pro Gly Ile Val His Phe Gly  
 20 25 30

Val Gly Asn Phe Phe Arg Ala His Glu Ala Phe Tyr Val Glu Gln Ile  
35 40 45

Leu Glu His Ala Pro Asp Trp Ala Ile Val Gly Val Gly Leu Thr Gly  
50 55 60

Ser Asp Arg Ser Lys Lys Lys Ala Glu Glu Phe Lys Ala Gln Asp Cys  
65 70 75 80

Leu Tyr Ser Leu Thr Glu Thr Ala Pro Ser Gly Lys Ser Thr Val Arg  
85 90 95

Val Met Gly Ala Leu Arg Asp Tyr Leu Leu Ala Pro Ala Asp Pro Glu  
100 105 110

Ala Val Leu Lys His Leu Val Asp Pro Ala Ile Arg Ile Val Ser Met  
115 120 125

Thr Ile Thr Glu Gly Gly Tyr Asn Ile Asn Glu Thr Thr Gly Ala Phe  
130 135 140

Asp Leu Glu Asn Ala Ala Val Lys Ala Asp Leu Lys Asn Pro Glu Lys  
145 150 155 160

Pro Ser Thr Val Phe Gly Tyr Val Val Glu Ala Leu Arg Arg Arg Trp  
165 170 175

Asp Ala Gly Gly Lys Ala Phe Thr Val Met Ser Cys Asp Asn Leu Arg  
180 185 190

His Asn Gly Asn Val Ala Arg Lys Ala Phe Leu Gly Tyr Ala Lys Ala  
195 200 205

Arg Asp Pro Glu Leu Ala Lys Trp Ile Glu Glu Asn Ala Thr Phe Pro  
210 215 220

Asn Gly Met Val Asp Arg Ile Thr Pro Thr Val Ser Ala Glu Ile Ala  
225 230 235 240

Lys Lys Leu Asn Ala Ala Ser Gly Leu Asp Asp Asp Leu Pro Leu Val  
245 250 255

Ala Glu Asp Phe His Gln Trp Val Leu Glu Asp Gln Phe Ala Asp Gly  
260 265 270

- Arg Pro Pro Leu Glu Lys Ala Gly Val Gln Met Val Gly Asp Val Thr  
275 280 285

Asp Trp Glu Tyr Val Lys Ile Arg Met Leu Asn Ala Gly His Val Met  
290 295 300

Leu Cys Phe Pro Gly Ile Leu Val Gly Tyr Glu Asn Val Asp Asp Ala  
305 310 315 320

012

Ile Glu Asp Ser Glu Leu Leu Gly Asn Leu Lys Asn Tyr Leu Asn Lys  
325 330 335

Asp Val Ile Pro Thr Leu Lys Ala Pro Ser Gly Met Thr Leu Glu Gly  
340 345 350

Tyr Arg Asp Ser Val Ile Ser Arg Phe Ser Asn Lys Ala Met Ser Asp  
355 360 365

Gln Thr Leu Arg Ile Ala Ser Asp Gly Cys Ser Lys Val Gln Val Phe  
370 375 380

Trp Thr Glu Thr Val Arg Arg Ala Ile Glu Asp Lys Arg Asp Leu Ser  
385 390 395 400

Arg Ile Ala Phe Gly Ile Ala Ser Tyr Leu Glu Met Leu Arg Gly Arg  
405 410 415

Asp Glu Lys Gly Gly Thr Tyr Glu Ser Ser Glu Pro Thr Tyr Gly Asp  
420 425 430

Ala Glu Trp Lys Leu Ala Lys Ala Asp Asp Phe Glu Ser Ser Leu Lys  
435 440 445



Leu Pro Ala Phe Asp Gly Trp Arg Asp Leu Asp Thr Ser Glu Leu Asp  
 450 455 460

Gln Lys Val Ile Val Leu Arg Lys Ile Ile Arg Glu Lys Gly Val Lys  
 465 470 475 480

Ala Ala Ile Pro Ala  
 485

<210> 3

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA

<400> 3

gctgctgagt gatccg

16

<210> 4

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA

<400> 4

gactgctact tcgatcc

17

<210> 5

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA

<400> 5

cctacaccta gcctgc

16

<210> 6

<211> 16

<212> DNA

<213> Artificial Sequence

*ai2*  
<220>

<223> synthetic DNA

<400> 6

cagtgccgtc atgagg

16

<210> 7

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA

<400> 7

tcctgatctc ggtgcg

16

<210> 8

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA

<400> 8

gatgcttcag cacggc

16

<210> 9

<211> 16

<212> DNA

<213> Artificial Sequence

*ar*

<220>

<223> synthetic DNA

<400> 9

gacgatcacg gaaggc

16

<210> 10

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA

<400> 10

ggttacgtgg tcgacg

16

<210> 11

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA

<400> 11

ctatacctga caggtcc

17

<210> 12

<211> 16

<212> DNA

<213> Artificial Sequence

*B12*  
<220>

<223> synthetic DNA

<400> 12

gcgcgatctg gatacg

16

<210> 13

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA

<400> 13

cgaggatctc gaacgg

16

<210> 14

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA

<400> 14

cggattgcta gcgatggc

18

<210> 15

<211> 47

<212> DNA

<213> Artificial Sequence

ai2  
<220>

<223> synthetic DNA

<400> 15

atcgaggatc ctcaatgatg atgatgatga tgggccggga tggcggc

47

<210> 16

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA

<400> 16

atcgaggatc cattcggctt ttagggtagc

30

<210> 17

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA

<400> 17

tagctgagct catgggacag atctgagc

28

Q12 <210> 18

<211> 10

<212> PRT

<213> Gluconobacter oxydans

<400> 18

Met	Ile	Thr	Arg	Glu	Thr	Leu	Lys	Ser	Leu
1				5					10

<210> 19

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA

<400> 19

taggaatatt tctcatgatt acgcgcgaaa ccc

33

<210> 20

<211> 16

<212> DNA

<213> Artificial Sequence

Q12

<220>

<223> synthetic DNA

<400> 20

gatgcttcag cacggc

16